

RAW SEQUENCE LISTING

**The Biotechnology Systems Branch of the Scientific and Technical
Information Center (STIC) no errors detected.**

Application Serial Number: 10/588,633
Source: JFWP
Date Processed by STIC: 08/15/2006

ENTERED

CRF Errors Edited by the STIC Systems Branch

Serial Number: 10/588,633

CRF Edit Date: 08/15/2006
Edited by: DA

___ Realigned nucleic acid/amino acid numbers/text in cases where the sequence text "wrapped" to the next line

___ Corrected the SEQ ID NO. Sequence numbers edited were:

___ Inserted or corrected a nucleic number at the end of a nucleic line. SEQ ID NO's edited:

~~___~~ Deleted: ___ invalid beginning/end-of-file text ; ___ page numbers

___ Inserted mandatory headings/numeric identifiers, specifically:

___ Moved responses to same line as heading/numeric identifier, specifically:

___ Other:



IFWP

RAW SEQUENCE LISTING

DATE: 08/15/2006

PATENT APPLICATION: US/10/588,633

TIME: 15:47:12

Input Set : N:\DA\pto.da.txt

Output Set: N:\CRF4\08112006\J588633.raw

3 <110> APPLICANT: VIEILLARD, Vincent
 4 DEBRE, Patrice
 6 <120> TITLE OF INVENTION: Polypeptide Derived from gp41, a Vaccine Composition
 7 Comprising Said Polypeptide, and Uses for Treating an
 8 Infection by an HIV virus in an Individual
 10 <130> FILE REFERENCE: CHEP:019US
 C--> 12 <140> CURRENT APPLICATION NUMBER: US/10/588,633
 13 <141> CURRENT FILING DATE: 2006-08-04
 15 <150> PRIOR APPLICATION NUMBER: PCT/EP2005/001395
 16 <151> PRIOR FILING DATE: 2005-02-07
 18 <150> PRIOR APPLICATION NUMBER: PCT/EP2004/01106
 19 <151> PRIOR FILING DATE: 2004-02-06
 21 <160> NUMBER OF SEQ ID NOS: 4
 23 <170> SOFTWARE: PatentIn Ver. 2.1
 25 <210> SEQ ID NO: 1
 26 <211> LENGTH: 1168
 27 <212> TYPE: PRT
 28 <213> ORGANISM: Homo sapiens
 30 <400> SEQUENCE: 1
 31 Met Ser Ile Val Ile Pro Leu Gly Val Asp Thr Ala Glu Thr Ser Tyr
 32 1 5 10 15
 34 Leu Glu Met Ala Ala Gly Ser Glu Pro Glu Ser Val Glu Ala Ser Pro
 35 20 25 30
 37 Val Val Val Glu Lys Ser Asn Ser Tyr Pro His Gln Leu Tyr Thr Ser
 38 35 40 45
 40 Ser Ser His His Ser His Ser Tyr Ile Gly Leu Pro Tyr Ala Asp His
 41 50 55 60
 43 Asn Tyr Gly Ala Arg Pro Pro Pro Thr Pro Pro Ala Ser Pro Pro Pro
 44 65 70 75 80
 46 Ser Val Leu Ile Ser Lys Asn Glu Val Gly Ile Phe Thr Thr Pro Asn
 47 85 90 95
 49 Phe Asp Glu Thr Ser Ser Ala Thr Thr Ile Ser Thr Ser Glu Asp Gly
 50 100 105 110
 52 Ser Tyr Gly Thr Asp Val Thr Arg Cys Ile Cys Gly Phe Thr His Asp
 53 115 120 125
 55 Asp Gly Tyr Met Ile Cys Cys Asp Lys Cys Ser Val Trp Gln His Ile
 56 130 135 140
 58 Asp Cys Met Gly Ile Asp Arg Gln His Ile Pro Asp Thr Tyr Leu Cys
 59 145 150 155 160
 61 Glu Arg Cys Gln Pro Arg Asn Leu Asp Lys Glu Arg Ala Val Leu Leu
 62 165 170 175
 64 Gln Arg Arg Lys Arg Glu Asn Met Ser Asp Gly Asp Thr Ser Ala Thr
 65 180 185 190

RAW SEQUENCE LISTING

DATE: 08/15/2006

PATENT APPLICATION: US/10/588,633

TIME: 15:47:12

Input Set : N:\DA\pto.da.txt

Output Set: N:\CRF4\08112006\J588633.raw

```

67 Glu Ser Gly Asp Glu Val Pro Val Glu Leu Tyr Thr Ala Phe Gln His
68      195      200      205
70 Thr Pro Thr Ser Ile Thr Leu Thr Ala Ser Arg Val Ser Lys Val Asn
71      210      215      220
73 Asp Lys Arg Arg Lys Lys Ser Gly Glu Lys Glu Gln His Ile Ser Lys
74 225      230      235      240
76 Cys Lys Lys Ala Phe Arg Glu Gly Ser Arg Lys Ser Ser Arg Val Lys
77      245      250      255
79 Gly Ser Ala Pro Glu Ile Asp Pro Ser Ser Asp Gly Ser Asn Phe Gly
80      260      265      270
82 Trp Glu Thr Lys Ile Lys Ala Trp Met Asp Arg Tyr Glu Glu Ala Asn
83      275      280      285
85 Asn Asn Gln Tyr Ser Glu Gly Val Gln Arg Glu Ala Gln Arg Ile Ala
86      290      295      300
88 Leu Arg Leu Gly Asn Gly Asn Asp Lys Lys Glu Met Asn Lys Ser Asp
89 305      310      315      320
91 Leu Asn Thr Asn Asn Leu Leu Phe Lys Pro Pro Val Glu Ser His Ile
92      325      330      335
94 Gln Lys Asn Lys Lys Ile Leu Lys Ser Ala Lys Asp Leu Pro Pro Asp
95      340      345      350
97 Ala Leu Ile Ile Glu Tyr Arg Gly Lys Phe Met Leu Arg Glu Gln Phe
98      355      360      365
100 Glu Ala Asn Gly Tyr Phe Phe Lys Arg Pro Tyr Pro Phe Val Leu Phe
101      370      375      380
103 Tyr Ser Lys Phe His Gly Leu Glu Met Cys Val Asp Ala Arg Thr Phe
104 385      390      395      400
106 Gly Asn Glu Ala Arg Phe Ile Arg Arg Ser Cys Thr Pro Asn Ala Glu
107      405      410      415
109 Val Arg His Glu Ile Gln Asp Gly Thr Ile His Leu Tyr Ile Tyr Ser
110      420      425      430
112 Ile His Ser Ile Pro Lys Gly Thr Glu Ile Thr Ile Ala Phe Asp Phe
113      435      440      445
115 Asp Tyr Gly Asn Cys Lys Tyr Lys Val Asp Cys Ala Cys Leu Lys Glu
116      450      455      460
118 Asn Pro Glu Cys Pro Val Leu Lys Arg Ser Ser Glu Ser Met Glu Asn
119 465      470      475      480
121 Ile Asn Ser Gly Tyr Glu Thr Arg Arg Lys Lys Gly Lys Lys Asp Glu
122      485      490      495
124 Asp Ile Ser Lys Glu Lys Asp Thr Gln Asn Gln Asn Ile Thr Leu Asp
125      500      505      510
127 Cys Glu Gly Ala Thr Asn Lys Met Lys Ser Pro Glu Thr Lys Gln Arg
128      515      520      525
130 Lys Leu Ser Pro Leu Arg Leu Ser Val Ser Asn Asn Gln Glu Pro Asp
131      530      535      540
133 Phe Ile Asp Asp Ile Glu Glu Lys Thr Pro Ile Ser Asn Glu Val Glu
134 545      550      555      560
136 Met Glu Ser Glu Glu Gln Ile Ala Glu Arg Lys Arg Lys Met Thr Arg
137      565      570      575
139 Glu Glu Arg Lys Met Glu Ala Ile Leu Gln Ala Phe Ala Arg Leu Glu

```

RAW SEQUENCE LISTING

DATE: 08/15/2006

PATENT APPLICATION: US/10/588,633

TIME: 15:47:12

Input Set : N:\DA\pto.da.txt

Output Set: N:\CRF4\08112006\J588633.raw

```

140          580          585          590
142 Lys Arg Glu Lys Arg Arg Glu Gln Ala Leu Glu Arg Ile Ser Thr Ala
143          595          600          605
145 Lys Thr Glu Val Lys Thr Glu Cys Lys Asp Thr Gln Ile Val Ser Asp
146          610          615          620
148 Ala Glu Val Ile Gln Glu Gln Ala Lys Glu Glu Asn Ala Ser Lys Pro
149 625          630          635          640
151 Thr Pro Ala Lys Val Asn Arg Thr Lys Gln Arg Lys Ser Phe Ser Arg
152          645          650          655
154 Ser Arg Thr His Ile Gly Gln Gln Arg Arg Arg His Arg Thr Val Ser
155          660          665          670
157 Met Cys Ser Asp Ile Gln Pro Ser Ser Pro Asp Ile Glu Val Thr Ser
158          675          680          685
160 Gln Gln Asn Asp Ile Glu Asn Thr Val Leu Thr Ile Glu Pro Glu Thr
161          690          695          700
163 Glu Thr Ala Leu Ala Glu Ile Ile Thr Glu Thr Glu Val Pro Ala Leu
164 705          710          715          720
166 Asn Lys Cys Pro Thr Lys Tyr Pro Lys Thr Lys Lys His Leu Val Asn
167          725          730          735
169 Glu Trp Leu Ser Glu Lys Asn Glu Lys Thr Gly Lys Pro Ser Asp Gly
170          740          745          750
172 Leu Ser Glu Arg Pro Leu Arg Ile Thr Thr Asp Pro Glu Val Leu Ala
173          755          760          765
175 Thr Gln Leu Asn Ser Leu Pro Gly Leu Thr Tyr Ser Pro His Val Tyr
176          770          775          780
178 Ser Thr Pro Lys His Tyr Ile Arg Phe Thr Ser Pro Phe Leu Ser Glu
179 785          790          795          800
181 Lys Arg Arg Arg Lys Glu Pro Thr Glu Asn Ile Ser Gly Ser Cys Lys
182          805          810          815
184 Lys Arg Trp Leu Lys Gln Ala Leu Glu Glu Glu Asn Ser Ala Ile Leu
185          820          825          830
187 His Arg Phe Asn Ser Pro Cys Gln Glu Arg Ser Arg Ser Pro Ala Val
188          835          840          845
190 Asn Gly Glu Asn Lys Ser Pro Leu Leu Leu Asn Asp Ser Cys Ser Leu
191          850          855          860
193 Pro Asp Leu Thr Thr Pro Leu Lys Lys Arg Arg Phe Tyr Gln Leu Leu
194 865          870          875          880
196 Asp Ser Val Tyr Ser Glu Thr Ser Thr Pro Thr Pro Ser Pro Tyr Ala
197          885          890          895
199 Thr Pro Thr His Thr Asp Ile Thr Pro Met Asp Pro Ser Phe Ala Thr
200          900          905          910
202 Pro Pro Arg Ile Lys Ser Asp Asp Glu Thr Cys Arg Asn Gly Tyr Lys
203          915          920          925
205 Pro Ile Tyr Ser Pro Val Thr Pro Val Thr Pro Gly Thr Pro Gly Asn
206          930          935          940
208 Thr Met His Phe Glu Asn Ile Ser Ser Pro Glu Ser Ser Pro Glu Ile
209 945          950          955          960
211 Lys Arg Arg Thr Tyr Ser Gln Glu Gly Tyr Asp Arg Ser Ser Thr Met
212          965          970          975

```

RAW SEQUENCE LISTING

DATE: 08/15/2006

PATENT APPLICATION: US/10/588,633

TIME: 15:47:12

Input Set : N:\DA\pto.da.txt

Output Set: N:\CRF4\08112006\J588633.raw

```

214 Leu Thr Leu Gly Pro Phe Arg Asn Ser Asn Leu Thr Glu Leu Gly Leu
215          980          985          990
217 Gln Glu Ile Lys Thr Ile Gly Tyr Thr Ser Pro Arg Ser Arg Thr Glu
218          995          1000          1005
220 Val Asn Arg Gln Cys Pro Gly Glu Lys Glu Pro Val Ser Asp Leu Gln
221      1010          1015          1020
223 Leu Gly Leu Asp Ala Val Glu Pro Thr Ala Leu His Lys Thr Leu Glu
224 1025          1030          1035          1040
226 Thr Pro Ala His Asp Arg Ala Glu Pro Asn Ser Gln Leu Asp Ser Thr
227          1045          1050          1055
229 His Ser Gly Arg Gly Thr Met Tyr Ser Ser Trp Val Lys Ser Pro Asp
230          1060          1065          1070
232 Arg Thr Gly Val Asn Phe Ser Val Asn Ser Asn Leu Arg Asp Leu Thr
233          1075          1080          1085
235 Pro Ser His Gln Leu Glu Val Gly Gly Gly Phe Arg Ile Ser Glu Ser
236      1090          1095          1100
238 Lys Cys Leu Met Gln Asp Asp Thr Arg Gly Met Phe Met Glu Thr Thr
239 1105          1110          1115          1120
241 Val Phe Cys Thr Ser Glu Asp Gly Leu Val Ser Gly Phe Gly Arg Thr
242          1125          1130          1135
244 Val Asn Asp Asn Leu Ile Asp Gly Asn Cys Thr Pro Gln Asn Pro Pro
245          1140          1145          1150
247 Gln Lys Lys Lys Ser Pro Val Gly Asn Phe Val Gly Ser Asn Val Val
248      1155          1160          1165
255 <210> SEQ ID NO: 2
256 <211> LENGTH: 262
257 <212> TYPE: PRT
258 <213> ORGANISM: Homo sapiens
260 <400> SEQUENCE: 2
261 Met Ala Trp Arg Ala Leu His His Trp Leu Leu Leu Leu Leu Phe Pro
262   1          5          10          15
264 Gly Ser Gln Ala Gln Ser Lys Ala Gln Val Leu Gln Ser Val Ala Gly
265          20          25          30
267 Gln Thr Leu Thr Val Arg Cys Gln Tyr Pro Pro Thr Gly Ser Leu Tyr
268          35          40          45
270 Glu Lys Lys Gly Trp Cys Lys Glu Ala Ser Ala Leu Val Cys Ile Arg
271          50          55          60
273 Leu Val Thr Ser Ser Lys Pro Arg Thr Met Ala Trp Thr Ser Arg Phe
274  65          70          75          80
276 Thr Ile Trp Asp Asp Pro Asp Ala Gly Phe Phe Thr Val Thr Met Thr
277          85          90          95
279 Asp Leu Arg Glu Glu Asp Ser Gly His Tyr Trp Cys Arg Ile Tyr Arg
280          100          105          110
282 Pro Ser Asp Asn Ser Val Ser Lys Ser Val Arg Phe Tyr Leu Val Val
283          115          120          125
285 Ser Pro Ala Ser Ala Ser Thr Gln Thr Pro Trp Thr Pro Arg Asp Leu
286      130          135          140
288 Val Ser Ser Gln Thr Gln Thr Gln Ser Cys Val Pro Pro Thr Ala Gly
289 145          150          155          160

```

RAW SEQUENCE LISTING

DATE: 08/15/2006

PATENT APPLICATION: US/10/588,633

TIME: 15:47:12

Input Set : N:\DA\pto.da.txt

Output Set: N:\CRF4\08112006\J588633.raw

```

291 Ala Arg Gln Ala Pro Glu Ser Pro Ser Thr Ile Pro Val Pro Ser Gln
292                165                170                175
294 Pro Gln Asn Ser Thr Leu Arg Pro Gly Pro Ala Ala Pro Ile Ala Leu
295                180                185                190
297 Val Pro Val Phe Cys Gly Leu Leu Val Ala Lys Ser Leu Val Leu Ser
298                195                200                205
300 Ala Leu Leu Val Trp Trp Gly Asp Ile Trp Trp Lys Thr Val Met Glu
301                210                215                220
303 Leu Arg Ser Leu Asp Thr Gln Lys Ala Thr Cys His Leu Gln Gln Val
304 225                230                235                240
306 Thr Asp Leu Pro Trp Thr Ser Val Ser Ser Pro Val Glu Arg Glu Ile
307                245                250                255
309 Leu Tyr His Thr Val Ala
310                260
314 <210> SEQ ID NO: 3
315 <211> LENGTH: 3507
316 <212> TYPE: DNA
317 <213> ORGANISM: Homo sapiens
319 <400> SEQUENCE: 3
320 atgagcatag tcatccattt ggggggtgat acagcagaga cgtcactt ggaaatggct 60
321 gcaggttcag aaccagaatc cgtagaagct agccctgtgg tagttgagaa atccaacagt 120
322 tatccccacc agttatatac cagcagctca catcattcac acagttacat tggtttgccc 180
323 tatgcggacc ataattatgg tgctcgctct cctccgacac ctccggcttc cctcctcca 240
324 tcagtcctta ttagcaaaaa tgaagtaggc atatttacca ctctaattt tgatgaaact 300
325 tccagtgtga ctacaatcag cacatctgag gatggaagtt atggtactga tgaaccagg 360
326 tgcatatgtg gttttacaca tgatgatgga tacatgatct gttgtgacaa atgcagcgtt 420
327 tggcaacata ttgactgcat ggggattgat aggcagcata ttctgatac atatctatgt 480
328 gaacgttgtc agcctaggaa tttggataaa gagagggcag tgctactaca acgccggaaa 540
329 agggaaaata tgcagatgg tgataccagt gcaactgaga gtggtgatga ggttcctgtg 600
330 gaattatata ctgcatttca gcatactcca acatcaatta ctttaactgc ttcaagagtt 660
331 tccaaagtta atgataaaag aaggaaaaaa agcggggaga aagaacaaca catttcaaaa 720
332 tgtaaaaagg catttcgtga aggatctagg aagtcacaa gagttaaggg ttcagctcca 780
333 gagattgatc cttcatctga tggttcaa at tttggatggg agacaaagat caaagcatgg 840
334 atggatcgat atgaagaagc aaataacaac cagtatagtg aggggtgttca gagggaggca 900
335 caaagaatag ctctgagatt aggcaatgga aatgacaaaa aagagatgaa taaatccgat 960
336 ttgaatacca acaatttgct cttcaaacct cctgtagaga gccatataca aaagaataag 1020
337 aaaattctta aatctgcaaa agatttgctt cctgatgcac ttatcattga atacagaggg 1080
338 aagtttatgc tgagagaaca gtttgaaagc aatgggtatt tctttaaaag accataacct 1140
339 tttgtgttat tctactctaa atttcatggg ctgaaatgt gtgttgatgc aaggactttt 1200
340 gggaaatgagg ctgattcat caggcggctt tgtacacca atgcagaggt gaggcagtaa 1260
341 attcaagatg gaaccataca tctttatat tattctatac acagtattcc aaagggaact 1320
342 gaaattacta ttgcctttga ttttgactat ggaaattgta agtacaaggt ggactgtgca 1380
343 tgcctcaaaag aaaacccaga gtgccctgtt ctaaaacgta gttctgaatc catggaaaaa 1440
344 atcaatagtg gttatgagac cagacggaaa aaaggaaaaa aagacgaaga tatttcaaaa 1500
345 gaaaaagata caaaaaatca gaataattact ttggattgtg aaggagcgac caacaaaatg 1560
346 aagagcccag aaactaaaca aagaaagctt tctccactga gactatcagt atcaaataat 1620
347 caggaaccag attttattga tgatatagaa gaaaaaactc ctattagtaa tgaagtagaa 1680
348 atggaatcag aggagcagat tgcagaaagg aaaaggaaga tgacaagaga agaaagaaaa 1740
349 atggaagcaa ttttgcaagc ttttgccaga cttgaaaaaa gagagaaaag aagagaacaa 1800

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/588,633

DATE: 08/15/2006

TIME: 15:47:13

Input Set : N:\DA\pto.da.txt

Output Set: N:\CRF4\08112006\J588633.raw

L:12 M:270 C: Current Application Number differs, Replaced Current Application Number



**Raw Sequence Listing before editing,
for reference only**



IFWP

RAW SEQUENCE LISTING

DATE: 08/14/2006

PATENT APPLICATION: US/10/588,633

TIME: 14:05:29

Input Set : N:\KEISHA\10588633.txt

Output Set: N:\CRF4\08142006\J588633.raw

3 <110> APPLICANT: VIEILLARD, Vincent

4 DEBRE, Patrice

6 <120> TITLE OF INVENTION: Polypeptide Derived from gp41, a Vaccine

Composition

7 Comprising Said Polypeptide, and Uses for Treating an

8 Infection by an HIV virus in an Individual

10 <130> FILE REFERENCE: CHEP:019US

C--> 12 <140> CURRENT APPLICATION NUMBER: US/10/588,633

13 <141> CURRENT FILING DATE: 2006-08-04

15 <150> PRIOR APPLICATION NUMBER: PCT/EP2005/001395

16 <151> PRIOR FILING DATE: 2005-02-07

18 <150> PRIOR APPLICATION NUMBER: PCT/EP2004/01106

19 <151> PRIOR FILING DATE: 2004-02-06

21 <160> NUMBER OF SEQ ID NOS: 4

23 <170> SOFTWARE: PatentIn Ver. 2.1

Does Not Comply
Corrected Diskette Needed

CP2-D

ERRORED SEQUENCES

381 <210> SEQ ID NO: 4

382 <211> LENGTH: 787

383 <212> TYPE: DNA

384 <213> ORGANISM: Homo sapiens

386 <400> SEQUENCE: 4

```

387 atggcctggc gaggcctaca ccactggcta ctgctgctgc tgttcccagg ctctcaggca 60
388 caatccaagg ctgaggtact tcaaagtgtg gcagggcaga cgctaaccgt gagatgccag 120
389 taccgcgcca cgggcagtct ctacgagaag aaaggctggt gtaaggaggc ttcagcactt 180
390 gtgtgcatca ggtagtcac cagctccaag ccaggacga tggcttgga cctctcgattc 240
391 acaatctggg acgaccctga tgctggcttc ttcactgtca ccatgactga tctaagagag 300
392 gaagactcag gacattactg gtgtagaatc taccgccctt ctgacaactc tgtctctaag 360
393 tccgtcagat tctatctggt ggtatctcca gcctctgcct ccacacagac cccctggact 420
394 ccccgcgacc tggctctctc acagaccag acccagagct gtgtgcctcc cactgcagga 480
395 gccagacaag cccctgagtc tccatctacc atccctgtcc cttctcagcc acagaactcc 540
396 acgctccgcc ctggccctgc agccccatt gccctggtgc ctgtgttctg tggactcctc 600
397 gtagccaaga gcctggtgct gtcagccctg ctgctctggt ggggggacat atggtggaaa 660
398 accgtgatgg agctcaggag cctggatacc caaaaagcca cctgccacct tcaacaggtc 720
399 acggaccttc cctggacctc agtttctca cctgtagaga gagaaatatt atatcacact 780
400 gttgcaa 787

```

E--> 403

1

deleted

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/588,633

DATE: 08/14/2006

TIME: 14:05:30

Input Set : N:\KEISHA\10588633.txt

Output Set: N:\CRF4\08142006\J588633.raw

L:12 M:270 C: Current Application Number differs, Replaced Current Application Number

L:403 M:254 E: No. of Bases conflict, this line has no nucleotides.